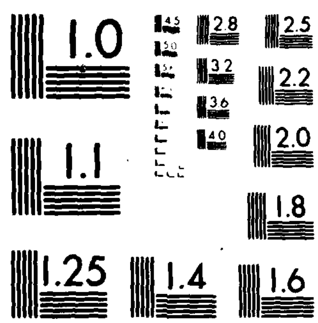
AD-A093 149 CALIFORNIA UNIV DAVIS INTERCOLLEGE DIV OF STATISTICS F/G 12/1
MAXIMUM LIKELIHOOD ESTIMATION FOR A DISCRETE MULTIVARIATE SHOCK--ETC(U)
AUG 80 R A BOYLES, F J SAMANIEGO AFOSR-77-3180
UNCLASSIFIED TR-21 AFOSR-TR-80-1158 NL

1 OF 1
AD A
093149

END
DATE
FILMED
1-81
DTIC



MICROCOPY RESOLUTION TEST CHART

AFOSR-TR- 80-1138

LEVEL

AD A093149

TECHNICAL REPORT SERIES
OF THE
INTERCOLLEGE DIVISION OF STATISTICS

DTIC
ELECTE
DEC 19 1980

University of California
Davis, California 95616

80 11 06 078

Maximum Likelihood Estimation for a Discrete Multivariate Shock Model*

Russell A. Boyles and Francisco J. Samaniego

Technical Report No. 21

August 1980

*This research was supported in part by the Air Force Office of Scientific Research under the grant AFOSR-77-3180.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered) UNCLASSIFIED

19 REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	18 AFOSR-TR-80-1158
2. GOVT ACCESSION NO.	AD-A093749
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle)	6 MAXIMUM LIKELIHOOD ESTIMATION FOR A DISCRETE MULTIVARIATE SHOCK MODEL,
5. TYPE OF REPORT & PERIOD COVERED	9 Interim rept.,
6. PERFORMING ORG. REPORT NUMBER	14 TR-21
7. AUTHOR(s)	10 Russell A. Boyles [redacted] Francisco J. Samaniego
8. CONTRACT OR GRANT NUMBER(s)	15 AFOSR-77-3180
9. PERFORMING ORGANIZATION NAME AND ADDRESS	University of California, Division of Statistics, Davis, California 95616
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS	16 61102F 2304/A5 17 A5
11. CONTROLLING OFFICE NAME AND ADDRESS	Air Force Office of Scientific Research/NM, Bolling AFB, Washington, DC 20332
12. REPORT DATE	11 Aug[redacted]980
13. NUMBER OF PAGES	20 12 33
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	
15. SECURITY CLASS. (of this report)	UNCLASSIFIED
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE	

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

maximum likelihood estimation, shock model, multivariate Bernoulli distribution

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A k-variate Bernoulli distribution with k+1 parameters is obtained as a shock model in which shocks are fatal to single components only or to all components simultaneously in a k-component system. The maximum likelihood estmator for model parameters if fully characterized. A simple iterative scheme is investigated, and it is shown that the scheme converges to the MLE for any seed in an interval whose endpoints depend only on the observed sample.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

412106

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

JOB

A k-variate Bernoulli distribution with $k+1$ parameters is obtained as a shock model in which shocks are fatal to single components only or to all components simultaneously in a k-component system. The maximum likelihood estimator for model parameters is fully characterized. A simple iterative scheme is investigated, and it is shown that the scheme converges to the MLE for any seed in an interval whose endpoints depend only on the observed sample.

I. INTRODUCTION

Let $Z_0, Z_1, \ldots, Z_k$ be independent Bernoulli variables, each with its own parameter p_i, $i=0,\ldots,k$. Let

$$Y_i = \min(Z_0, Z_i) \qquad i=1,\ldots,k,$$

and consider the distribution of the random vector $\underset{\sim}{Y} = (Y_1,\ldots,Y_k)$. This distribution is a $(k+1)$-parameter submodel of the multivariate Bernoulli distributions studied in Boyles and Samaniego (1980). This paper is dedicated to maximum likelihood estimation for this submodel, henceforth to be denoted by MVB$(k+1)$.

The model MVB$(k+1)$ can be motivated as follows: Suppose a k component system may be subjected to two kinds of shocks during an observation period of length T_0. A shock of type 1 is fatal to a single component and has no effect on other components. Define

$$Z_i = \begin{cases} 0 & \text{if shock fatal to component } i \text{ occurs by time } T_0 \\ 1 & \text{otherwise.} \end{cases}$$

A shock of type 2 is simultaneously fatal to all components. Define

$$Z_0 = \begin{cases} 0 & \text{if universal shock occurs by time } T_0 \\ 1 & \text{otherwise.} \end{cases}$$

and define

$$p_i = P(Z_i = 1), \qquad i=0,1,\ldots,k.$$

Then $Y_i = 1$ if and only if no shock fatal to component i occurs by time T_0. The model above may be viewed as a discrete analogue of the submodel of Marshall and Olkin's (1967) multivariate exponential (MVE) distribution with single and universal shocks only. This submodel of the MVE distribution has been studied extensively by Proschan and Sullo (1976).

The general multivariate Bernoulli model studied in Boyles and Samaniego (1980) postulates the existence of $2^k - 1$ shocks, each selectively fatal to a particular subset of the k components of the system. Maximum likelihood estimation for the general model poses substantial analytical difficulties, and is an unsolved and perhaps intractable problem. In Boyles and Samaniego (1980), the invariance property of MLE's is used to produce an asymptotically optimal estimator that is in fact equal to the MLE with limiting probability one. The submodel considered in this paper is the only MVB shock model we have examined in which the MLE itself can be fully characterized. The MVB$(k+1)$ distribution thus has the following characteristics: (1) It is a model for random vectors with positively dependent components, (2) its parameter space is of relatively low dimension ($k+1$ instead of $2^k - 1$ for the general model), permitting efficient estimation with moderate sample sizes, (3) it is a reasonable model in experiments where the primary or only cause of simultaneous failure of components is a catastrophic or universal shock, (4) maximum likelihood estimation is fully tractable. Because of these characteristics, the authors feel that the model merits the separate detailed study presented in this paper.

In Section II, we characterize the MLE of the parameters of MVB($k+1$), showing that in the most complex case, the MLE is a simple function of the smallest root of a certain k^{th} degree polynomial. In Section III, we give a simple iterative scheme which converges quickly to the desired root.

II. MAXIMUM LIKELIHOOD ESTIMATION

Suppose a sample of size n is taken from MVB($k+1$). The likelihood function for the sample is given by

$$L = p_0^T Q_0^{n-T} \prod_{i=1}^{k} p_i^{N_i} (1 - p_i)^{T-N_i} \tag{2.1}$$

where

$$Q_0 = P\{\underset{\sim}{Y} = \underset{\sim}{0}\} = 1 - p_0\{1 - \prod_{i=1}^{k}(1 - p_i)\},$$

$$T = \# \text{ times } \underset{\sim}{Y} \neq \underset{\sim}{0}$$

$$N_i = \# \text{ times } Y_i = 1 \qquad (i=1,2,\ldots,k).$$

Note that

$$\max\{N_1,\ldots,N_k\} \leq T \leq \sum_{i=1}^{k} N_i. \tag{2.2}$$

If $T=0$ then $N_i=0\ \forall i$ so (2.1) becomes $L=Q_0^n$ which is maximized by $Q_0 = 1$. Since

$$Q_0 = 1 - p_0\{1 - \prod_{i=1}^{k}(1-p_i)\} \tag{2.3}$$

we maximize L in this case by taking either

$$\hat{p}_0 = 0,\ \hat{p}_i \text{ arbitrary} \quad (i=1,\dots,k)$$

or

$$\hat{p}_0 \text{ arbitrary},\ \hat{p}_1 = \hat{p}_2 = \dots = \hat{p}_k = 0.$$

Now assume $0 < T = N_1$ and $N_2 = N_3 = \dots = N_k = 0$. (2.1) becomes

$$L = p_0^{N_1} Q_0^{n-N_1} p_0^{N_1} \prod_{i=2}^{k}(1-p_i)^{N_1}. \tag{2.4}$$

It is evident from (2.3) that, for any fixed values of p_0 and p_1, Q_0 is maximized by setting $p_2 = p_3 = \dots = p_k = 0$. Thus (2.4) is maximized for fixed p_0 and p_1 by setting $p_2 = \dots = p_k = 0$. We are left with

$$L = p_0^{N_1}\, Q_0^{n-N_1}\, p_1^{N_1} = (p_0p_1)^{N_1}(1 - p_0p_1)^{n-N_1},$$

so that in this case (2.1) is maximized by any choice of $(\hat{p}_0,\hat{p}_1,\dots,\hat{p}_k) \in [0,1]^{k+1}$ satisfying

$$\hat{p}_0\hat{p}_1 = \frac{1}{n} N_1$$

$$\hat{p}_2 = \dots = \hat{p}_k = 0.$$

Suppose now that $T > 0$, $N_1 > 0,\dots,N_\ell > 0$, $N_{\ell+1} = \dots = N_k = 0$ where $2 \le \ell \le k$. For any fixed $p_0,p_1,\dots,p_\ell$ we maximize Q_0 by setting $p_{\ell+1} = \dots = p_k = 0$. In this case

$$L = p_0^T Q_0^{n-T}\{\prod_{i=1}^{\ell} p_i^{N_i}(1-p_i)^{T-N_i}\} \prod_{i=\ell+1}^{k}(1-p_i)^T,$$

so we maximize L (for fixed $p_0, p_1, \ldots, p_\ell$) by setting $p_{\ell+1} = \cdots = p_k = 0$. Having done this we are left with the task of maximizing (2.1) subject to $k = \ell \geq 2$, $T > 0$, $N_1 > 0, \ldots, N_\ell > 0$. Thus, without loss of generality, we assume $k \geq 2$, $T > 0$, $N_1 > 0$, $N_2 > 0, \ldots, N_k > 0$ in the remainder of this paper.

From (2.1) it can be shown that the likelihood equations

$$\frac{\partial \ln L}{\partial p_i} = 0 \qquad (i=0,1,\ldots,k)$$

are equivalent to the system

$$\begin{aligned} \mu_0 &= p_0\{1 - \prod_{i=1}^{k}(1-p_i)\} \\ \mu_1 &= p_1 p_0 \\ &\vdots \\ \mu_k &= p_k p_0 \end{aligned} \tag{2.5}$$

where $\mu_0 = n^{-1}T$ and $\mu_i = n^{-1}N_i$ $(i=1,\ldots,k)$. Solutions of (2.5) are in one-to-one correspondence with solutions of

$$\mu_0 = p_0\{1 - \prod_{i=1}^{k}(1-\mu_i/p_0)\}.$$

Setting $x = p_0^{-1}$ we see that solutions of (2.5) are in one-to-one correspondence with solutions of

$$1 - \mu_0 x = \prod_{i=1}^{k}(1-\mu_i x). \tag{2.6}$$

Note that, because of (2.2) and our standing assumptions ($N_i > 0 \;\forall\, i$) we have

$$\min\{1, \sum_{i=1}^{k} \mu_i\} \geq \mu_0 \geq \mu_{(k)} = \max\{\mu_1, \ldots, \mu_k\}$$

and (2.7)

$$\mu_{(1)} = \min\{\mu_1, \ldots, \mu_k\} > 0.$$

<u>Lemma</u>. The system (2.5) has a solution $\hat{\underset{\sim}{p}} = (\hat{p}_0, \hat{p}_1, \ldots, \hat{p}_k) \in [0,1]^{k+1}$ iff (2.6) has a solution $\hat{x} \in [1, \mu_0^{-1}]$. Moreover, $\hat{\underset{\sim}{p}} \in (0,1)^{k+1}$ iff $\hat{x} \in (1, \mu_0^{-1})$.

<u>Proof</u>. We already know that solutions of (2.5) are in one-to-one correspondence with solutions of (2.6). Assume $\hat{\underset{\sim}{p}} \in [0,1]^{k+1}$. Then

$$\mu_0 = \hat{p}_0\{1 - \prod_{i=1}^{k} (1-\hat{p}_i)\} \leq \hat{p}_0 \leq 1$$

so $\hat{x} = \hat{p}_0^{-1} \in [1, \mu_0^{-1}]$ and solves (2.6). Conversely, if $\hat{x} \in [1, \mu_0^{-1}]$ then $\hat{p}_0 = \hat{x}^{-1} \in [\mu_0, 1] \subset [0,1]$ and, since

$$0 < \mu_i \leq \mu_0 \leq \hat{p}_0,$$

we have $\hat{p}_i = \mu_i \hat{p}_0^{-1} \in [0,1]$ for $i=1,\ldots,k$.

The verification of the second statement is the same, but we will include it for completeness. Assume $\hat{\underset{\sim}{p}} \in (0,1)^{k+1}$. Then $\prod_{i=1}^{k} (1-\hat{p}_i) > 0$ so

$$\mu_0 = \hat{p}_0\{1 - \prod_{i=1}^{k} (1-\hat{p}_i)\} < \hat{p}_0 < 1$$

and $\hat{x} = \hat{p}_0^{-1} \in (1,\mu_0^{-1})$. Conversely, if $\hat{x} \in (1,\mu_0^{-1})$ then $\hat{p}_0 = \hat{x}^{-1} \in (\mu_0,1) \subset (0,1)$ and, since

$$0 < \mu_i \le \mu_0 < \hat{p}_0,$$

we have $\hat{p}_i = \mu_i \hat{p}_0^{-1} \in (0,1)$. This completes the proof of the lemma.

Define

$$h(x) = \prod_{i=1}^{k} (1 - \mu_i x). \tag{2.8}$$

h is a k^{th} degree polynomial with roots $\mu_1^{-1},\ldots,\mu_k^{-1}$. Moreover

$$h(0) = 1,\ h(\mu_0^{-1}) = \prod_{i=1}^{k}\left(1 - \frac{\mu_i}{\mu_0}\right) \ge 0,$$

and $h > 0$ on $(-\infty,\mu_{(k)}^{-1})$. For $x < \mu_{(k)}^{-1}$ we have

$$h'(x) = -\sum_{i=1}^{k} \mu_i \prod_{j \ne i} (1 - \mu_j x) < 0$$

and

$$h''(x) = \sum_{i=1}^{k} \sum_{j \ne i} \mu_i \mu_j \prod_{\substack{\ell \ne j \\ \ell \ne i}} (1 - \mu_\ell x) > 0.$$

In particular $h'(0) = -\sum_{i=1}^{k} \mu_i \le -\mu_0$ = slope of line $y = 1 - \mu_0 x$. This line and the curve $y = h(x)$ intersect at $x = 0$. If $h'(0) = -\mu_0$ we have $h(x) > 1 - \mu_0 x$ for all $x \in (0,\mu_{(k)}^{-1}]$. (In this case $h'' > 0$ forces $\mu_{(k)}^{-1} > \mu_0^{-1}$.) If $h'(0) < -\mu_0$ the line $y = 1 - \mu_0 x$ and the curve $y = h(x)$

intersect at exactly one point $\hat{x} \in (0, \mu_{(k)}^{-1}]$. In fact, we have $\hat{x} \in (0, \mu_0^{-1}]$, since $1 - \mu_0 x < 0$ for $x > \mu_0^{-1}$ while $h > 0$ on $(-\infty, \mu_{(k)}^{-1})$.

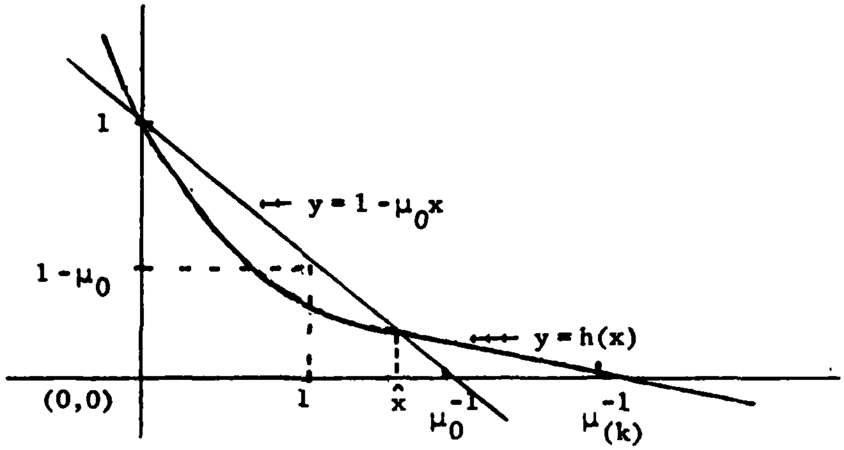


Figure I

In spite of the picture, we are not excluding the possibility that $\mu_0^{-1} = \mu_{(k)}^{-1}$, or that $\hat{x} \leq 1$. In fact, we have

$$\hat{x} \geq 1 \quad \text{iff} \quad \prod_{i=1}^{k} (1 - \mu_i) = h(1) \leq 1 - \mu_0 . \tag{2.9}$$

Assume now that $\prod_{i=1}^{k} (1 - \mu_i) > 1 - \mu_0$, i.e., $\hat{x} \in [0,1)$. By the lemma, the likelihood equations have no root in $[0,1]^{k+1}$, so we seek to maximize L over the boundary of $[0,1]^{k+1}$. Now since $N_i > 0 \ \forall\, i$ and $T > 0$, we see from (2.1) that $L = 0$ on the faces $\{p_i = 0\}$ for $i = 0,1,\ldots,k$. On the other hand, since

$$1 - \mu_j > \prod_{i=1}^{k} (1 - \mu_i) > 1 - \mu_0$$

we have $\mu_j < \mu_0 \Rightarrow T - N_j > 0$ $(j=1,\ldots,k)$. Thus $L = 0$ on the faces $\{p_i = 1\}$, $i=1,\ldots,k$. On $\{p_0 = 1\}$ we have

$$L = \left\{\prod_{i=1}^{k}(1-p_i)\right\}^{n-T} \prod_{i=1}^{k} p_i^{N_i}(1-p_i)^{T-N_i}$$

$$= \prod_{i=1}^{k} p_i^{N_i}(1-p_i)^{n-N_i} .$$

Thus the MLE in this case is given by

$$\hat{p}_0 = 1, \ \hat{p}_i = n^{-1}N_i \qquad (i=1,\ldots,k). \tag{2.10}$$

Now assume $\prod_{i=1}^{k}(1-\mu_i) \le 1-\mu_0$, i.e., $\hat{x} \in [1,\mu_0^{-1}]$. We will show that that the MLE in this case is given by

$$\hat{p}_0 = \hat{x}^{-1}, \ \hat{p}_i = \mu_i\hat{x} \qquad (i=1,\ldots,k). \tag{2.11}$$

We first consider the likelihood function (2.1) as a function $\tilde{L}$ of $x \in [1,\mu_0^{-1}]$, i.e., we examine its values along the curve

$$\{(x^{-1},\mu_1 x,\ldots,\mu_k x) : 1 \le x \le \mu_0^{-1}\}$$

in $[0,1]^{k+1}$. First let $x \in (1,\mu_0^{-1})$. Then we have

$$\tilde{L}(x) = x^{-T}\left\{1-x^{-1}\left[1 - \prod_{i=1}^{k}(1-\mu_i x)\right]\right\}^{n-T}$$

$$\cdot \prod_{i=1}^{k}(\mu_i x)^{N_i}(1-\mu_i x)^{T-N_i}$$

$$= x^{-n}\left\{\frac{x-1+\prod_{i=1}^{k}(1-\mu_i x)}{\prod_{i=1}^{k}(1-\mu_i x)}\right\}^{n-T} \prod_{i=1}^{k}(\mu_i x)^{N_i}(1-\mu_i x)^{n-N_i}$$

$$= \left\{x^{-1}\left[\frac{x-1+h(x)}{h(x)}\right]^{1-\mu_0} \prod_{i=1}^{k}(\mu_i x)^{\mu_i}(1-\mu_i x)^{1-\mu_i}\right\}^{n} .$$

Now

$$\ell(x) = n^{-1}\ln \tilde{L}(x)$$

$$= -\ln x + (1-\mu_0)\ln[x-1+h(x)]$$

$$- (1-\mu_0)\ln h(x) + \sum_{i=1}^{k}\mu_i \ln\left[\frac{\mu_i x}{1-\mu_i x}\right] + \sum_{i=1}^{k}\ln(1-\mu_i x)$$

$$= -\ln x + (1-\mu_0)\ln[x-1+h(x)]$$

$$+ \mu_0 \ln h(x) + \sum_{i=1}^{k}\mu_i \ln\left[\frac{\mu_i x}{1-\mu_i x}\right].$$

Differentiating, we obtain

$$\ell'(x) = -x^{-1} + \frac{(1-\mu_0)[1+h'(x)]}{x-1+h(x)} + \frac{\mu_0 h'(x)}{h(x)}$$

$$+ x^{-1}\sum_{i=1}^{k}\left[\frac{\mu_i}{1-\mu_i x}\right].$$

Since

$$h'(x) = -\sum_{i=1}^{k}\mu_i \prod_{j\neq i}(1-\mu_j x)$$

$$= -h(x)\sum_{i=1}^{k}\left[\frac{\mu_i}{1-\mu_i x}\right]$$

we have

$$\ell'(x) = -x^{-1} + \frac{(1-\mu_0)[1+h'(x)]}{x-1+h(x)} + [\mu_0 - x^{-1}]\frac{h'(x)}{h(x)}$$

$$= -\frac{h(x)+(1-\mu_0 x)h'(x)}{xh(x)} + \frac{(1-\mu_0)[1+h'(x)]}{x-1+h(x)}$$

$$= [(1-\mu_0 x) - h(x)]\left[\frac{h(x)-h'(x)(x-1)}{xh(x)[x-1+h(x)]}\right].$$

Since $h(x) > 0$, $h'(x) < 0$, and $x-1 > 0$ for $x \in (1,\mu_0^{-1})$ we see that

$$\ell'(x) \text{ is } \begin{Bmatrix} > 0 \\ = 0 \\ < 0 \end{Bmatrix} \text{ iff } \begin{Bmatrix} h(x) < 1-\mu_0 x \\ h(x) = 1-\mu_0 x \\ h(x) > 1-\mu_0 x \end{Bmatrix}.$$

If $\hat{x} \in (1, \mu_0^{-1})$ this means

$$\ell'(x) \text{ is } \begin{Bmatrix} > 0 \\ = 0 \\ < 0 \end{Bmatrix} \text{ iff } \begin{Bmatrix} 1 < x < \hat{x} \\ x = \hat{x} \\ \hat{x} < x < \mu_0^{-1} \end{Bmatrix}.$$

If $\hat{x} = 1$, we have

$$\ell'(x) < 0 \qquad \forall\ x \in (1, \mu_0^{-1}).$$

If $\hat{x} = \mu_0^{-1}$, we have

$$\ell'(x) > 0 \qquad \forall\ x \in (1, \mu_0^{-1}).$$

Since $\ell(x)$ is continuous, we have shown that

$$\ell(\hat{x}) = \sup_{x \in [1, \mu_0^{-1}]} \ell(x), \quad \ell(\hat{x}) > \ell(x) \quad \text{for } x \neq \hat{x}.$$

Consequently,

$$\tilde{L}(\hat{x}) = \sup_{x \in [1, \mu_0^{-1}]} \tilde{L}(x), \quad \tilde{L}(\hat{x}) > \tilde{L}(\hat{x}) \quad \text{for } x \neq \hat{x}.$$

Now consider L on the boundary of $[0,1]^{k+1}$. Since $T, N_i > 0 \ \forall\ i$ $L = 0$ on $\{p_i = 0\}$, $i=0,1,\ldots,k$. On $\{p_0 = 1\}$ we know from (2.10) that the maximum value of L is given by

$$\prod_{i=1}^{k} (n^{-1}N_i)^{N_i} (1 - n^{-1}N_i)^{n-N_i} = \tilde{L}(1).$$

Now consider the face $\{p_j = 1\}$ for some $j=1,\ldots,k$. If $T - N_j > 0$ then $L = 0$ on $\{p_j = 1\}$. If $T - N_j = 0$, then

$$L = p_0^{\,T}(1-p_0)^{n-T}\prod_{\substack{i=1\\ i\neq j}}^{k} p_i^{\,N_i}(1-p_i)^{T-N_i},$$

so the maximum value is obtained by evaluating L at

$$\hat{p}_0 = n^{-1}T, \quad \hat{p}_i = \frac{N_i}{T} \qquad (i=1,\ldots,k) \tag{2.12}$$

(since $\frac{N_j}{T} = 1$). The maximum value is

$$\begin{aligned} L(\hat{p}_0,\hat{p}_1,\ldots,\hat{p}_k) &= L\left(n^{-1}T, \frac{N_1}{T},\ldots,\frac{N_k}{T}\right) \\ &= L\left(\mu_0, \frac{\mu_1}{\mu_0},\ldots,\frac{\mu_k}{\mu_0}\right) \\ &= \tilde{L}(\mu_0^{-1}). \end{aligned}$$

If $\hat{x} \in (1, \mu_0^{-1})$ the likelihood equations have a unique root $\hat{\underset{\sim}{p}} \in (0,1)^{k+1}$, where $\hat{\underset{\sim}{p}}$ is given by (2.11). Since

$$L(\hat{\underset{\sim}{p}}) = \tilde{L}(\hat{x}) > \max\{\tilde{L}(1), \tilde{L}(\mu_0^{-1})\}$$

L achieves its maximum in $(0,1)^{k+1}$. By uniqueness, $\hat{\underset{\sim}{p}}$ is the MLE.

If $\hat{x} = 1$, L has no root in $(0,1)^{k+1}$. If $\hat{\underset{\sim}{p}}$ is given by (2.11),

$$L(\hat{\underset{\sim}{p}}) = \tilde{L}(\hat{x}) = \tilde{L}(1) > \tilde{L}(\mu_0^{-1}).$$

Thus $\hat{\underset{\sim}{p}}$ maximizes L on the boundary of $[0,1]^{k+1}$, so $\hat{\underset{\sim}{p}}$ is the MLE.

If $\hat{x} = \mu_0^{-1}$, L has no root in $(0,1)^{k+1}$. If $\hat{\underset{\sim}{p}}$ is given by (2.11)

$$L(\hat{\underset{\sim}{p}}) = \tilde{L}(\hat{x}) = \tilde{L}(\mu_0^{-1}) > \tilde{L}(1).$$

Thus $\hat{\underset{\sim}{p}}$ maximizes L on the boundary of $[0,1]^{k+1}$, so $\hat{\underset{\sim}{p}}$ is the MLE.

This completes the proof that $\hat{\underset{\sim}{p}}$ given by (2.11) is the MLE in the case $\prod_{i=1}^{k}(1-\mu_i) \le 1-\mu_0$.

The following table gives the complete description of the MLE $\hat{\underset{\sim}{p}}$ of the MVB(k+1) parameter $\underset{\sim}{p}$.

<table>
<tr><td colspan="3">$T = 0$</td><td>$\hat{p}_0 = 0$, $\hat{p}_i$ arbitrary $(i=1,\dots,k)$
or
$\hat{p}_0$ arbitrary, $\hat{p}_1 = \dots = \hat{p}_k = 0$</td></tr>
<tr><td rowspan="3">$T > 0$</td><td colspan="2">$N_j = T$, $N_i = 0$ $(i \neq j)$</td><td>$\hat{p}_0\hat{p}_j = n^{-1}N_j$, $\hat{p}_i = 0$ $(i \neq j)$</td></tr>
<tr><td rowspan="2">$N_1 > 0,\dots,N_\ell > 0, N_{\ell+1} = \dots = N_k = 0$
$(2 \le \ell \le k)$</td><td>$\prod_{i=1}^{\ell}(1-\mu_i) > 1-\mu_0$
$(\mu_i = n^{-1}N_i, \mu_0 = n^{-1}T)$</td><td>$\hat{p}_0 = 1$, $\hat{p}_i = n^{-1}N_i$ $(i=1,\dots,k)$</td></tr>
<tr><td>$\prod_{i=1}^{\ell}(1-\mu_i) \le 1-\mu_0$
$(\mu_i = n^{-1}N_i, \mu_0 = n^{-1}T)$</td><td>$\hat{p}_0 = \hat{x}^{-1}$, $\hat{p}_i = x\, n^{-1}N_i$ $(i=1,\dots,k)$
where $\hat{x}$ is the unique solution in $[1,\mu_0^{-1}]$ of $1-\mu_0 x = \prod_{i=1}^{\ell}(1-\mu_i x)$</td></tr>
</table>

III. AN ITERATIVE SCHEME CONVERGING TO THE MLE

In the following discussion, we exclude those data configuations for which $\hat{\underset{\sim}{p}}$ can be computed explicitly. In other words, we assume the following:

$$\begin{aligned} &\text{(i)} \quad T > 0,\ N_1 > 0, \ldots, N_k > 0 \\ &\text{(ii)} \quad \prod_{i=1}^{k}(1-\mu_i) < 1-\mu_0 \\ &\text{(iii)} \quad T > \max\{N_1, \ldots, N_k\}. \end{aligned} \tag{3.1}$$

We have shown (but perhaps not stated explicitly) that under (i) and (ii) the equation

$$1-\mu_0 x = \prod_{i=1}^{k}(1-\mu_i x) \tag{3.2}$$

has a unique solution $\hat{x} \in (1, \mu_0^{-1}]$. If, in addition, we assume (iii), then

$$\prod_{i=1}^{k}(1-\mu_i\mu_0^{-1}) > 0$$

since $\mu_0 > \mu_i$ $(i=1,2,\ldots,k)$. Thus, (3.2) is not satisfied by $x = \mu_0^{-1}$, so $\hat{x} \in (1, \mu_0^{-1})$. On the other hand, if (iii) fails then we have $T - N_j = 0 \Longrightarrow 1-\mu_j\mu_0^{-1} = 0$ for some j, so (3.2) is satisfied by $x = \mu_0^{-1}$, hence $\hat{x} = \mu_0^{-1}$ by uniqueness.

The result of these considerations is that (3.1) is equivalent to $\hat{x} \in (1,\mu_0^{-1})$, hence the situation described in (3.1) is precisely that in which we have no explicit formula for $\hat{x}$, hence no explicit formula for $\hat{\underset{\sim}{p}}$. We now give a simple iterative solution of (3.2) under conditions (3.1).

Let $\hat{x}^{(0)} \in [1,\mu_0^{-1}]$ and define

$$\begin{aligned}\hat{x}^{(m)} &= \mu_0^{-1}\{1 - \prod_{i=1}^{k}(1-\mu_i\hat{x}^{(m-1)})\}\\ &= \mu_0^{-1}\{1 - h(\hat{x}^{(m-1)})\} \qquad (3.3)\end{aligned}$$

for $m = 1,2,3,\ldots$. Figure II below indicates graphically how the iterative scheme proceeds.

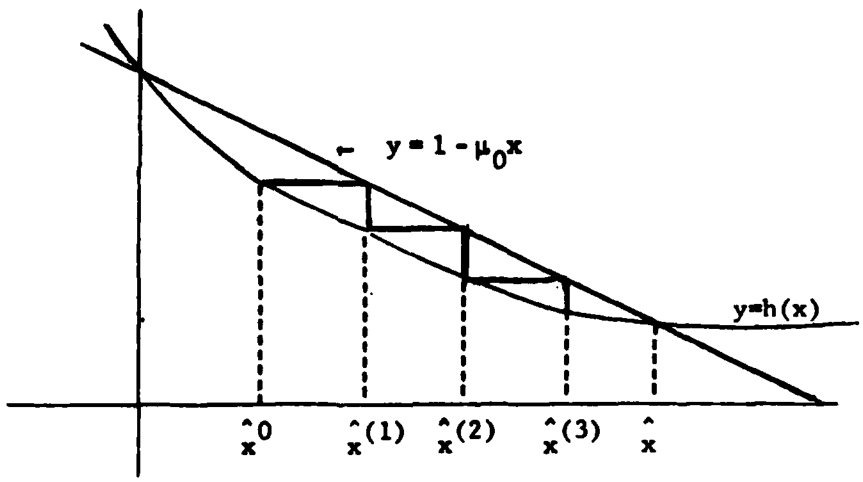


Figure II

<u>Theorem</u>. $\hat{x}^{(m)} \to \hat{x}$ as $m \to \infty$.

<u>Proof</u>. If $\hat{x}^{(0)} = \hat{x}$ then $\hat{x}^{(m)} = \hat{x}$ $\forall$ m and we are done. Assume $\hat{x}^{(0)} < \hat{x}$. Recall that

$$h(x) = \prod_{i=1}^{k}(1-\mu_i x) \qquad (1 \le x \le \mu_0^{-1})$$

and that $h' < 0$ on $[1,\mu_0^{-1}]$. Since $\hat{x}^{(0)} < \hat{x}$ we have, by the definition of $\hat{x}^{(1)}$, that

$$1-\mu_0\hat{x}^{(0)} > h(\hat{x}^{(0)}) = 1-\mu_0\hat{x}^{(1)}.$$

Moreover, by definition of $\hat{x}$ we have

$$1-\mu_0\hat{x}^{(1)} = h(\hat{x}^{(0)}) > h(\hat{x}) = 1-\mu_0\hat{x} .$$

Hence

$$\hat{x}^{(0)} < \hat{x}^{(1)} < \hat{x} .$$

Now assume

$$\hat{x}^{(0)} < \hat{x}^{(1)} < \ldots < \hat{x}^{(m)} < \hat{x} . \tag{3.4}$$

By (3.4) and the definition of $\hat{x}^{(m+1)}$, we have

$$1-\mu_0\hat{x}^{(m)} > h(\hat{x}^{(m)}) = 1-\mu_0\hat{x}^{(m+1)}.$$

Moreover, by definition of $\hat{x}$ we have

$$1-\mu_0\hat{x}^{(m+1)} = h(\hat{x}^{(m)}) > h(\hat{x}) = 1-\mu_o\hat{x} .$$

Hence

$$\hat{x}^{(m)} < \hat{x}^{(m+1)} < \hat{x} ,$$

which proves (3.4) with m replaced by $m+1$. We have now proven by induction that $\{\hat{x}^{(m)} : m = 0,1,\ldots\}$ is a bounded increasing sequence. Let

$$y = \lim_{m\to\infty} \hat{x}^{(m)}.$$

Then, by the continuity of h, we have

$$y = \lim_{m\to\infty} \hat{x}^{(m)}$$

$$= \lim_{m\to\infty} \mu_0^{-1}\{1 - h(\hat{x}^{(m-1)})\}$$

$$= \mu_0^{-1}\{1 - h(y)\}$$

or

$$1 - \mu_0 y = \prod_{i=1}^{k} (1 - \mu_i y).$$

Thus $y = \hat{x}$, as required.

Now assume $\hat{x}^{(0)} > \hat{x}$. The proof is essentially the same as in the preceding case, but we include it for completeness. We have

$$1 - \mu_0 \hat{x}^{(0)} < h(\hat{x}^{(0)}) = 1 - \mu_0 \hat{x}^{(1)}$$

and

$$1 - \mu_0 \hat{x}^{(1)} = h(\hat{x}^{(0)}) < h(\hat{x}) = 1 - \mu_0 \hat{x},$$

hence

$$\hat{x} < \hat{x}^{(1)} < \hat{x}^{(0)}.$$

Now assume

$$\hat{x} < \hat{x}^{(m)} < \ldots < \hat{x}^{(1)} < \hat{x}^{(0)} .$$

Then

$$1 - \mu_0 \hat{x}^{(m)} < h(\hat{x}^{(m)}) = 1 - \mu_0 \hat{x}^{(m+1)}$$

and

$$1 - \mu_0 \hat{x}^{(m+1)} = h(\hat{x}^{(m)}) < h(\hat{x}) = 1 - \mu_0 \hat{x} ,$$

hence

$$\hat{x} < \hat{x}^{(m+1)} < \hat{x}^{(m)} < \ldots < \hat{x}^{(1)} < \hat{x}^{(0)} .$$

This proves by induction that $\{\hat{x}^{(m)}\}$ is decreasing and bounded below. Let $y = \lim_{m\to\infty} \hat{x}^{(m)}$. Then we show as before that $y = \hat{x}$. This completes the proof.

For any fixed n, define the estimates $\hat{\underset{\sim}{p}}^{(m)}$ $(m=0,1,2,\ldots)$ by

$$\hat{\underset{\sim}{p}}^{(m)} = \begin{cases} \hat{\underset{\sim}{p}} & \text{if (3.1) does not hold} \\ ([\hat{x}^{(m)}]^{-1}, \mu_1 \hat{x}^{(m)}, \ldots, \mu_k \hat{x}^{(m)}) & \text{if (3.1) holds.} \end{cases}$$

<u>Corollary</u>. For any fixed n, $\hat{\underset{\sim}{p}}^{(m)} \to \hat{\underset{\sim}{p}}$ a.s. $(m \to \infty)$.

The proof consists of quoting the preceding theorem.

Two likely candidates for $\hat{x}^{(0)}$ are 1 and μ_0^{-1}. It would be useful to develop some criteria for deciding which $\hat{x}^{(0)}$ to use. Also, the first iterate $\hat{\underset{\sim}{p}}^{(1)}$ might be worthy of study.

REFERENCES

Boyles, R.A. and Samaniego, F.J. (1980). Modeling and Inference for Positively Dependent Random Variables in Dichotomous Experiments. Technical Report No. 19, Division of Statistics, University of California, Davis.

Marshall, A. and Olkin, I. (1967). A Multivariate Exponential Distribution. Journal of the American Statistical Association, **62**, 30-44.

Proschan, F. and Sullo, P. (1976). Estimating the Parameters of a Multivariate Exponential Distribution. Journal of the American Statistical Association, **71**, 465-72.